

## REMARKS

In the outstanding Office Action, claims 28 to 38 were presented for examination. Claims 30-34, 36 and 38 were rejected on the basis of 35 U.S.C. §103 as being unpatentable in view of references to Minematsu, Shpiro and Lessac.

The Office Action has been most carefully studied. In this amendment applicant has canceled claims 30-34, 36 and 38, without prejudice, and has added new claims 40-59. The new claims have been carefully written to avoid any questions under 35 U.S.C. §112, in accordance with the guidelines and requirements set forth in the outstanding Office Action. Accordingly, as will be discussed in detail below, it is believed that the application is clearly in condition for allowance.

### *New Method Claim 40*

New claim 40 is directed to the invention as it relates to a computerized speech recognition method which employs, as basic processing units, not only individual words but also more extended utterances, namely phrases. As defined in new claim 40, the inventive speech recognition method employs a language model database having a phrase model database in addition to the conventional word model database. Each of the phrase and word model databases comprises digital representations, and associated graphic representations, of properly pronounced spoken phrases and words, respectively. In the claimed method, an input speech signal is converted into digital representations of audible sound components of the input speech. The digital representation of an audible sound component of the input speech is then compared to the digital representations of spoken words and spoken phrases in the word and phrase model databases to determine a match with one of the word or phrase digital representations. The graphic representation associated with the matched word or phrase is then output as a recognized speech component.

The invention as now claimed in claim 40 helps overcome problems of system hardware capacity and speed of performance and may also enhance accuracy by

employing phrases as processing units in addition to words. Utilizing a phrase model database, a given phrase present in the database can be recognized in the input speech in one matching step, and output as one graphic element, for example, an alphanumeric element, depending upon the language. For the first time the claimed invention provides a speech recognition method which can recognize a phrase in speech input to a computerized system and output a corresponding graphic in a single iteration of the speech recognition cycle. For the phrase, the sequential processing of multiple words is avoided, enabling improvements in processing efficiency.

New Claim 40 is fully supported by the disclosure as may be seen for example from Fig. 1, noting step 28 (top center) "User Input Words and Phrases", step 44 (bottom right) "Error Type Phrase" and step 58 (top right) "Exercise Phrase Student Input" as well as block 22 (center) "Exercise Phrase Model" and block 26 (center) "Exercise Phrase Error Model", and from the detailed description which is replete with disclosure regarding the use of a phrase model database and phrase recognition.

Applicant's usage of the term "phrase" in describing the invention is exemplified at page 27, line 19, where the usage of "phrase" is described as referencing, *inter alia*, at least parts from two different words or two or more words. It will be understood that the phrases represented and recognized in the claimed invention may also comprise more extended utterances.

***Claimed Invention Not Suggested by Combination of Minematsu and Shpiro***

The foregoing discussion and the explanations already of record are believed to make clear that the invention as now claimed in independent base claims 40 and 57 is clearly and patentably distinguished from the art of record, whether considered alone or in combination. In particular, the claimed invention is believed not remotely suggested by Minematsu or Shpiro, any combination of the two references or any other art known to applicant.

Minematsu discloses a speech recognition apparatus and method for correctly recognizing an English word from a non-native English pronunciation (abstract). As is made clear throughout the specification, Minematsu's unit of recognition is the word not the phrase. Minematsu correlates an analogous word (mispronounced word) with a candidate word (properly pronounced word). Where a plurality of words is erroneously recognized as a another plurality of words, Minematsu still correlates the analogous word to the candidate word, albeit in this case ("The Fourth Method" in column 12, lines 39-42), "a plurality of words" is included with each of the analogous word and the candidate word. Minematsu does not suggest a phrase model database comprising digital and associated graphic representations of phrases. Nor does Minematsu employ phrase matching as required by applicant's claims wherein a digital representation of an input speech sound component is compared with the digital representation of a spoken phrase.

The Minematsu method is simply a method for identifying repetitive individual letter, phoneme, and word mispronunciations by a person who is a native speaker of another language and who is attempting to speak a second language.

Shpiro discloses an interactive language training apparatus, rather than a speech recognition method. Shpiro employs two phrase databases, Databases E and F (column 4, lines 39-46. However, Shpiro's phrase databases are not components of language model used for speech recognition, they are not suitable for speech recognition and they are not used by Shpiro for speech recognition as it is understood in the art, i.e. the identification and rendering in text of unknown speech.

Nothing about Shpiro's disclosure suggests database E or F could constitute a language model database providing a full range of common words and phrases in a language adequate for speech recognition. The user input to Shpiro's system comprises "expected audio responses" not "unknown speech" as called for by applicant's claims. Furthermore, as described at column 4, lines 61-67, Shpro's method measures the

likelihood of a match between the user's actual audio response and the expected response, i.e. Shpiro is attempting to measure how well the user has understood the language and responded in the language. The language may be understood from column 1, lines 17-24, to be a foreign language from the user's perspective. This is quite different from speech recognition wherein unknown speech, typically, although not necessarily from a native speaker, is recognized and output as text. Nothing in Shpiro suggests how the problem of improving processing efficiency and accuracy of speech recognition methods such as Minematsu's can be solved.

Accordingly, even a combination of Shpiro with Minematsu would not suggest applicant's invention as now claimed.

#### *New System Claim 57*

New base claim 57 relates to a computerized system for performing a speech recognition method and is directed to similar subject matter to the subject matter of claim 40. Specifically, the claimed system has a speech recognition database which includes a phrase model database as is defined in the claim. Accordingly, claim 57 is believed allowable for the reasons claim 40 is believed allowable, as set forth above.

#### *Dependent Claims*

Claims 41-56 and 58-59 depend from base claims 40 and 57, respectively, and are therefore believed allowable with claims 40 and 57 for the reasons that claims 40 and 57 are believed allowable. Dependent claims 41-56 and 58-59 are furthermore believed clearly and patentably distinguished from the art of record, and therefore allowable, by the additional meaningful limitations they recite.

More particularly, Claim 42 specifically recites that employing one or more speech-trained speakers to input words and phrases into the word and phrase model databases, which is not remotely suggested by Minematsu, Shpiro, any combination of Minematsu and Shpiro, or any other art of record or known to applicant.

Furthermore, Claim 43 specifically recites that the word model database and/or the phrase model database comprise audio recordings, optionally in high quality format, to provide for audible playback of the respective words and/or phrases, which is not remotely suggested by Minematsu, Shpiro, any combination of Minematsu and Shpiro, or any other art of record or known to applicant. Applicant has discovered that computer errors may result in misrecognition of a particular error, mistaken acceptance of a mispronunciation, or mistaken rejection of a proper pronunciation. The claimed audio recordings of database words and/or phrases may be used to check or verify system performance in this respect.

Alternative limbs of claim 44 which depends from claim 42, call for error models of word and/or phrase or both word and phrase mispronunciations and for these mispronunciations optionally to be spoken by a speaker who normally makes such mispronunciations. The art does not remotely suggest employment of phrase error models. Nor does the art suggest the clear distinctions and recognition benefits which can be obtained by the claimed combination of trained speakers inputting correct pronunciations and "normal" mispronouncers inputting mispronunciations.

In addition, claim 46 requires that the word model database and/or the phrase model database comprise a collection of word and/or phrase models respectively, associated with one or more mispronunciations. This is because if a person makes a speech pronunciation error of a particular type, it is likely that the same speaker makes certain other errors which have common characteristics with other pronunciation errors in the group; which is neither suggested nor addressed by the art.

Moreover, Claim 48 in its optional alternative calls for speech training in accordance with the Lessac method. Nothing in the art suggests that Lessac system human voice training techniques can be beneficially employed in a computerized speech recognition method. Nor does the art suggest that Lessac system techniques might be

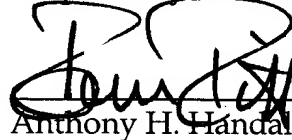
employed in a computerized interactive training method to enable users to effectively correct mispronunciations.

Claim 53 provides exercise word and/or phrase models which pursuant to claim 54 can be associated in groups having common mispronunciation characteristics; which features are also not remotely suggested by Minematsu, Shpiro, any combination of Minematsu and Shpiro, or any other art of record or known to applicant.

In view of the above amendments and the discussion relating thereto, it is respectfully submitted that the instant application, is in condition for allowance. Such action is most earnestly solicited. If for any reason the Examiner feels that consultation with Applicant's attorney would be helpful in the advancement of the prosecution, he is invited to call the telephone number below for an interview.

Respectfully submitted,

By:



Anthony H. Handal

Reg. No. 26,275 Ph: (212) 536-4870

Roger Pitt

Reg. No. 46,996 Ph: (212) 536-4867

KIRKPATRICK & LOCKHART LLP  
599 Lexington Avenue (32nd Floor)  
New York, NY 10022-6030

**Certificate of Mailing**

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope, postage prepaid, addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on November 24, 2004.



Anthony H. Handal Reg. No. 26,275

Roger Pitt Reg. No. 46,996